Viewpoints

2006 annual report for

ROSES

Applied Information Systems Research Program

Creon Levit[1] and Paul Gazis[2]

[1]NASA Ames Research Center, Advanced Supercomputing Division
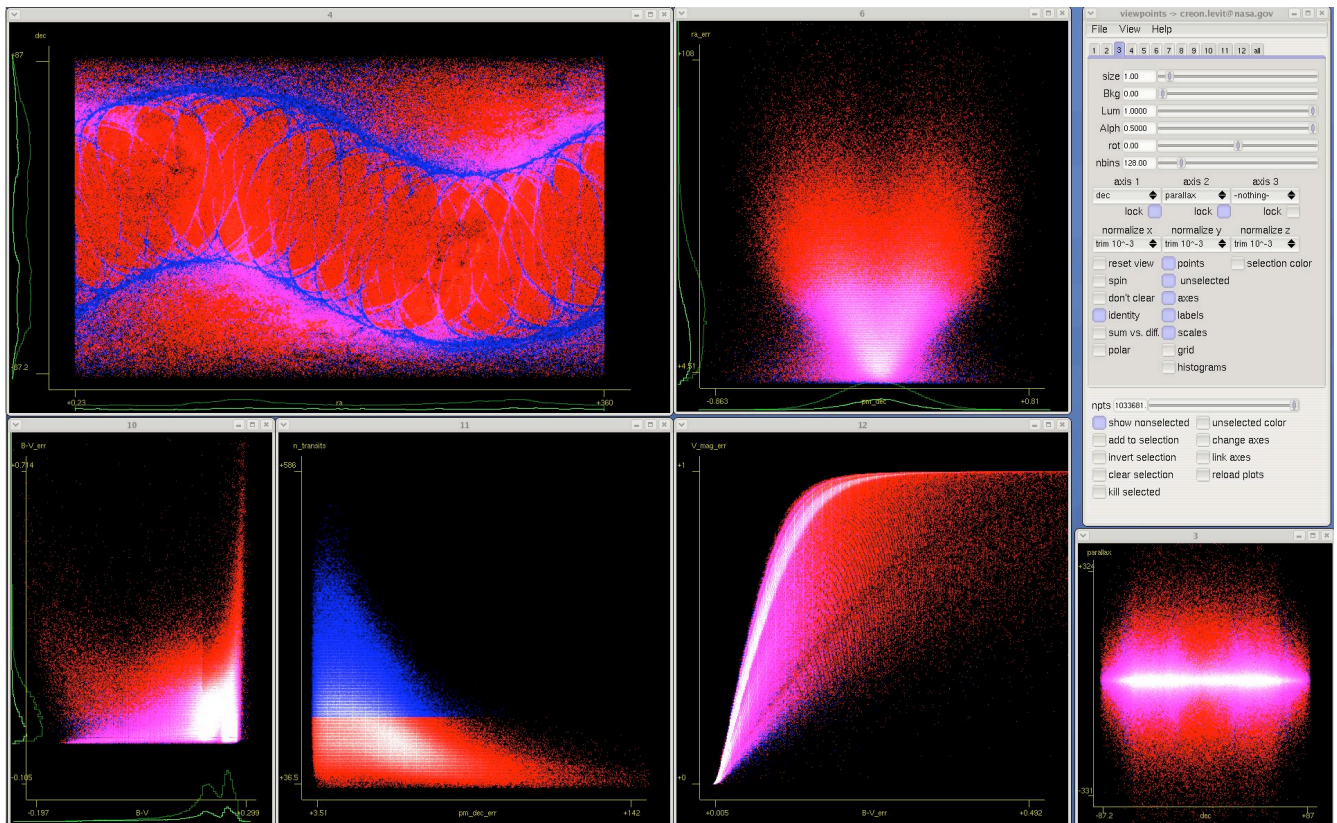[2]SETI Inst., NASA Ames Research Center, Space Sciences Division

figure 1 – Interactive visual analysis more than a million stars with 21 parameters per star (from the Hipparcos mission) using viewpoints running on a laptop

Accomplishments in 2006:

1. Viewpoints is now used by many scientists and engineers both inside and outside of NASA. It has been downloaded by over 60 users, about half of whom are from various NASA centers. Some example of viewpoints' current uses are:

    a. Project Constellation (CEV & CLV - design trade studies, postprocessing of CFD data from the Columbia supercomputer, abort analysis, trajectory studies)

    b. NASA Engineering and Safety Center (Stardust sample return capsule trajectory reconstruction, aircraft safety data visualization)

    c. Vehicle systems health monitoring (space shuttle main engine fault prediction)

    d. Aeronautics (Mars flyer simulations, multiparameter wing optimization)

    e. Astrophysics (galaxy surveys, stellar astrophysics, gamma ray astrophysics, many others)

    f. Space sciences (heliospheric physics, planetary science)

    g. Earth science (analysis of MODIS, GOES, and AVRIS and data)

    h. Computation finance (statistical arbitrage, portfolio management)

    i. Biology (influenza virus analysis and vaccine design)

2. Viewpoints now runs native on all major technical computing platforms: Linux, Apple OSX (both PowerPC and Intel), and Microsoft Windows. Its operation, interface, source code, and impressive performance are the same on all platforms.

3. Viewpoints is now directly callable from matlab.

4. There is a viewpoints website (astrophysics.arc.nasa.gov/viewpoints) containing an introduction to the program, screenshots, and downloadable (binary) distributions for all platforms. A viewpoints wiki is under construction and almost ready to go live.

5. Viewpoints presentations and/or demonstrations have been made at the following conferences, workshops, and facilities:

NASA Ames code S&T ROSES workshop	Ames Research Center	Creon Levit	3/08/06
Apple Computer science and technology group forum	Apple Computer	Creon Levit	2/16/06
NASA Ames Discovery and Systems Health group (demo)	Ames Research Center	Creon Levit	4/12/06
NASA Ames Discovery and Systems Health group (seminar)	Ames Research Center	Creon Levit	5/10/06
American Geophysical Union spring joint assembly	San Francisco, CA	Paul Gazis	5/23/06 - 5/26/06
Interface 2006: 38th symp. on the interface of statistics, computing & applications	Pasadena, CA	Creon Levit	5/24/06 - 5/27/06
208th Meeting of the American Astronomical Society	Calgary, Alberta Canada	Paul Gazis	6/4/06 - 6/8/06

NESC workshop on Data Mining Applications in Aeronautics and Space Exploration	NASA Ames	Creon Levit	6/20/06 - 6/21/06
Crew Launch Vehicle simulation group	NASA Ames	Creon Levit	8/08/06
John's Hopkins University	Baltimore, MD	Michael Way	04/05/06
NASA Headquarters	Washington, DC	Michael Way	04/05/06
Niels Bohr Institute	Copenhagen, Denmark	Michael Way	06/15/06
University of Austria	Vienna, Austria	Michael Way	06/06/06
The American Natural History Museum	New York, NY	Michael Way	04/07/06
U of Missouri	St. Louis, MO	Michael Way	11/18/05
Princeton University	Princeton, NJ	Michael Way	04/06/06

6. Viewpoints' functionality and user interface have been significantly enhanced to include:

 a. the ability to create and destroy plot windows on the fly.

 b. the ability to read or reread ascii or binary input files, and the ability to write binary data files on the fly.

 c. support for much larger datasets and more dimensions without recompilation – limited only by available (virtual) memory.

 d. the ability to change parameters "globally" across all plots at once, or "locally" on a single plot.

 e. the ability to "lock" (and unlock) individual plots so they are immune from (susceptible to) global operations.

 f. support for multiple selections in multiple colors.

 g. rudimentary semi-automated browsing through high dimensional data.

 h. display of marginal histograms, with bin counts that are independently adjustable in real time.

 i. support for "forward search" - successive applications of outlier detection, outlier removal, and renormalization.

 j. display of three dimensional scatterplots with real time rotation, scaling, etc.

 k. real-time display of quantitative data for individual axes as well as for the current selection.

 l. graceful error recovery when attempting to read corrupt and ill-formatted input files.

 m. the ability to be run from the command line, or from an icon.

 n. on line help.

7. Viewpoints is now being actively developed by multiple programmers:

 a. The source code is managed using "subversion" - a state of the art distributed revision control system (see subversion.tigris.org).

b. We have initiated the process of releasing viewpoints as an open-source project on sourceforge under the NASA Open Source Agreement (see opensource.arc.nasa.gov).

Most importantly, people like viewpoints; It is easy to use, and it does what it does *very* well. The authors' favorite anecdote: One normally quite reserved NASA scientist who leads a large data analysis group praised it, saying "viewpoints is like a work of art".

Plans for 2007 (pending support for 2007):

Our original proposal described a second year (2007) development plan combining increased functionality (in user interface, graphics, and data management) with increased performance. Now, with an additional year of experience and user input, we have prioritized and refined this development plan. The new development plan includes the following, prioritized separately within each category:

1. Data management and user interface:
    a. the ability to save the program's complete state to a configuration file at any time.
    b. the ability to restore the program's state from a configuration file.
    c. the ability or initialize the program's state from a configuration file specified at startup.
    d. the ability to write ascii data files along with user-provided header information in user-specified formats.
    e. improvements to the MATLAB interface.
    f. development of a basic IDL interface.
    g. improvements to the GUI, including an enhanced file-chooser and provisions to handle the configuration files described above.
    h. improvements to the help window.
    i. an error / warning / logging window.
2. graphics and visualization
    a. real-time display of additional quantitative information about selections and/or the ability to specify selections quantitatively
    b. utilization of (textured) point sprites for rendering large-sized points and a variety of other symbols
    c. automatic inclusion of index and random index columns to support simple subsampling.
    d. support for specifying and rendering data as vectors, timeseries, spectra, and parallel coordinates
    e. better control of 2D overplotting, e.g. based on selections.
    f. arbitrary 2D rotation of plots.
    g. better 3D view manipulation controls, e.g. virtual trackball.
    h. simpler control of selections, selection colors, and editing of selections.
    i. more capable grids and axis tickmarks (real-time, of course)
    j. gradient (variable color, or "soft") selections, thus allowing points to be colored based on

any variable, as well as on combinations of variables.

    k. multidimensional (D>3) plots and grand tours.

3. machine learning

    a. more advanced (i.e. multi-dimensional) outlier detection.

    b. better control of histograms, including support for the latest methods of determining (initial) bin locations.

    c. mutual information and correlation functionals between all pairs and triples of variables will be computed in a background thread. Scatterplot array layout can be ordered based on those functionals.

    d. interactive projection pursuit functionals

    e. density derivatives estimates

4. performance enhancements

    a. rendering from openGL vertex buffer objects.

    b. fast selections using presorted data.

    c. an openGL 2.0 vertex program to render vertex arrays where individual vertex components come from separate arrays. This will conserve memory (GPU) memory and allow for display of larger datasets.

Some Testimonials from current users:

"Viewpoints is a powerful, fast, and efficient method of doing data analysis on very large data files. We have used the software extensively to analyze data from a variety of sources, including earth science and space science data. The interface is intuitive and the system's response time is very fast. We use the software on both linux and windows, and have found that the system works well in both operating systems. We believe that the integration between viewpoints and matlab will be very useful for increasing the speed and efficiency with which we address NASA applications."

    Ashok N. Srivastava, Ph.D.
    Intelligent Data Understanding Group Lead
    NASA Ames Research Center

"Viewpoints is really very much what we've been looking for browsing aerodynamic datasets in a quick way. I'm very jazzed about having it available, and available on the desktop."

    Mike Aftosmis
    Algorithms and Applications Branch, Advanced Supercomputing Davison
    NASA Ames Research Center

"My work is on aerodynamic shape optimization. I've tested viewpoints on my Mars Flyer dataset and also a small dataset based on a CEV-like capsule. It works well for me."

> Marian Nemec, PhD.
> ELORET Corp., NASA Ames Research Center

"I have been using viewpoints since the first day I learned of it on a variety of different datasets. For example, I recently used it to challenge some referee comments on an Astrophysical Journal publication. The relevant analysis would have taken me several days to respond to with conventional plotting software, but with viewpoints I was able to do it in a matter of minutes!

"The interface is not bloated with a large number of useless features (like so many other graphical plotting programs) and it is quite fast with even large datasets. I've used it for everything from looking for patterns in datasets with millions of points and a few attributes per point to hundreds of thousands of points with 100s of attributes. I have also shown it to several colleagues in the past year and each time they appeared astounded by its abilities when demonstrated.

"It's a truly unique piece of software that has changed the way I work."

> Michael Way, PhD.
> Astrophysics Branch, Space Sciences Division
> NASA Ames Research Center

Conclusion

The AISR program's call for proposals to which we responded was for three year grants. Our viewpoints proposal was selected and awarded one year preliminary "seed funding" with the following summary of reviewers' evaluations:

> "…. The concept is sound and innovative, but *there are reservations concerning the plan and approach to accomplish.* The proposal is relevant and good fit with AISR objectives, and has strong enough appeal for NASA science endeavors to warrant seed funding to refine the proposal and develop a stronger plan and proof of concept." (emphasis added).

We think we have addressed these reviewers reservations fully during FY 2006: We have accomplished all of our first year technical objectives and more. We have a significant, growing user community who are successfully and enthusiastically applying viewpoints to problems and missions across the agency. And we have refined our second year development plan to incorporate the priorities of our users while still maintaining a strong applied research focus.

Viewoints feels like it might, possibly, become a "killer app" for science and engineering. We hope ROSES and the AISR program will consider awarding us an additional year or two of funding so we can continue to develop, distribute, support, and apply viewpoints to the benefit of an ever-growing set of NASA users, programs, and missions.